



Available online at : <http://bit.ly/InfoTekJar>

## InfoTekJar : Jurnal Nasional Informatika dan Teknologi Jaringan

ISSN (Print) 2540-7597 | ISSN (Online) 2540-7600



# Analisis Kelayakan Kredit Berbasis Algoritma K-Nearest Neighbor (Studi Kasus: Koperasi AKU)

Rusda Wajhillah<sup>1</sup>, Irsyad Hafizh Ubaidallah<sup>1</sup>, Saeful Bahri<sup>2</sup>

<sup>1</sup> Universitas Bina Sarana Informatika, Sukabumi, 43121, Indonesia

<sup>2</sup> STMIK Nusa Mandiri, Jakarta, 43111, Indonesia

## KEYWORDS

Koperasi, K-NN, Kredit

## CORRESPONDENCE

Phone: +62 (8111) 189808

E-mail: [rusda.rwh@bsi.ac.id](mailto:rusda.rwh@bsi.ac.id)

## ABSTRACT

Kredit macet merupakan sumber ancaman terbesar dalam menjalankan operasi bisnis, dan mampu menyebabkan kesulitan keuangan pada masa mendatang, dan telah tercatat 40% kredit macet pada sejumlah koperasi di Salatiga, 10 milyar kredit macet di Kabupaten Bantul, serta tumpukan kredit macet yang rata-rata mencapai 10% menyebabkan peningkatan koperasi tidak aktif hingga 29%. Analisa kredit sangat dibutuhkan untuk menekan resiko terjadinya kredit bermasalah. Salah satu cara untuk mengurangi resiko kredit bermasalah yaitu menggunakan sebuah algoritma yang mampu mengklasifikasikan data kelayakan pemberian kredit, salah satu algoritma yang telah sukses diterapkan oleh beberapa peneliti yaitu algoritma *K-Nearest Neighbor* hal ini disebabkan karena algoritma ini cocok digunakan untuk dataset yang berukuran tidak terlalu besar seperti pada dataset penelitian ini yang hanya terdapat 276 data yang dibagi kedalam 2 kategori 94 macet 182 lancar. Dari dataset yang di uji hasil yang didapatkan akurasi tertinggi sebesar 79,45% pada nilai K=1, dengan rata-rata akurasi 73,696%, dan nilai AUC tertinggi didapat pada K=9 dengan nilai sebesar 0,811, dengan nilai akurasi tersebut algoritma K-NN memiliki kinerja cukup baik dalam proses klasifikasi data kelayakan pemberian kredit.

## INTRODUCTION

Kredit merupakan penyediaan uang berdasarkan pinjam meminjam yang mewajibkan pihak peminjam melunasi hutangnya setelah jangka waktu tertentu dengan penambahan bunga [1]. Sedangkan koperasi merupakan badan hukum yang memisahkan kekayaan para anggotanya sebagai modal usaha, dan dapat memberikan pinjaman kepada anggota, serta mewajibkan anggota untuk melunasi dalam jangka waktu tertentu dan membayar jasa.

Selain sumber penghasilan, kredit merupakan sumber ancaman terbesar dalam operasi bisnis [2]. Selain itu, kredit yang bermasalah juga berpotensi menyebabkan kesulitan keuangan dimasa mendatang. Dalam penelitian yang dilakukan oleh Wahyono & Cahyono (2015) tercatat 40% kredit macet pada sejumlah koperasi di Salatiga, 10 milyar kredit macet di Kabupaten Bantul, serta tumpukan kredit macet yang rata-rata mencapai 10% dalam

Kredit Usaha Rakyat yang menyebabkan meningkatnya koperasi tidak aktif hingga 29%.

Terdapat beberapa penelitian mengenai analisa kredit yang dibuat oleh peneliti seperti Adi dan Winarko pada tahun 2015 [2] mengenai analisa tingkat keamanan kredit menggunakan model Naïve Bayes [3].

mengenai analisa permohonan kredit dengan menggunakan model decision tree C4.5.

K-Nearest Neighbor (KNN) merupakan algoritma sederhana yang bekerja dengan menitik beratkan jarak antara data training dengan data testing [4], selain itu KNN juga cocok untuk digunakan dalam dataset yang tidak terlalu besar [5].

Berdasarkan penjelasan diatas, penelitian ini kami menggunakan algoritma K-Nearest Neighbor untuk menganalisa kredit, yang diharapkan mampu menyelesaikan permasalahan tentang klasifikasi kredit yang bermasalah.

## RELEATEAD RESEARCH

Penelitian yang dilakukan oleh Adi & Winarko pada tahun 2015 mengenai prediksi tingkat keamanan pemberian kredit dengan menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine*. Dalam penelitian ini mendapatkan hasil 80% keakuratan *Naïve Bayes*, dan 72% keakuratan *Support Vector Machine* untuk masing – masing menggunakan 100 data sampel. Penelitian yang dilakukan oleh Amin, Indwiarti, & Sibaroni pada tahun 2015 mengenai keputusan permohonan kredit dengan menggunakan algoritma C4.5, dalam penelitian ini mendapatkan hasil keakuratan 74,5%.

Penelitian yang dilakukan oleh Setiawan dan Dayanti [3] mengenai klasifikasi citra retina menggunakan *K-Nearest Neighbor* untuk mendeteksi makulopati diabetik (kelainan pada mata akibat rusaknya pembuluh darah akibat komplikasi penyakit diabetes mellitus), dalam penelitian ini mendapatkan hasil keakuratan 93,33%.

Penelitian yang dilakukan oleh Budianita et al. tahun 2015 [4] mengenai implementasi pengolahan citra dan klasifikasi *K-Nearest Neighbor* yang digunakan untuk membangun aplikasi pembeda daging sapi dengan daging babi, dalam penelitian ini mendapatkan hasil keakuratan hingga 88,75%.

## METHOD

Metode penelitian yang digunakan merupakan metode experiment dengan menggunakan model CRISP-DM, yang meliputi 6 tahapan berikut tahapan yang dimaksud:

### Business Understanding

fase business understanding ini difokuskan untuk mencari tujuan dari penelitian. Dalam penelitian ini akan difokuskan dalam menganalisa kelayakan kredit dengan menggunakan K-Nearest Neighbor sebagai algoritma klasifikasi

### Data Understanding

Collecting, analizing, and evaluation quality of data. The data will be used in this research is secondary data obtained from Koperasi AKU. Terdapat 327 data, dengan sepuluh atribut diantaranya total kredit macet sebanyak 111 data, dan kredit lancar sebanyak 216 data, sedangkan untuk atribut yang terdapat pada data tersebut adalah No (nomor keanggotaan koperasi), Nama Ketua (Nama ketua kelompok usaha), Nama Kelompok (nama asosiasi kelompok UPPKS), Jumlah Anggota (Jumlah anggota asosiasi kelompok), Jenis Usaha (Jenis usaha yang dijalankan), Simpanan Pokok, Simpanan Wajib, Simpanan Sukarela, Jumlah Simpanan, dan Pinjaman, serta Keterangan sebagai label

### Data Preparation

Data yang ada akan dilakan pembersihan dengan cara membuang data yang duplikat, data yang kosong dan melakukan seleksi terhadap atribut yang tidak berpengaruh terhadap hasil analisa, setelah dilakukan pembersihan data maka didapat sebanyak 276 data dengan sebaran data sebanyak 94 data yang didentifikasi macet, dan 182 data yang teridentifikasi lancar.

Proses pemodelan menggunakan K-NN mengharuskan data yang diuji kedalam model harus memiliki class atau target berupa numerik [8], lebih jelas dapat dilihat pada tabel 1 data dan hasil pembersihan.

Tabel 1. Data hasil pembersihan

Member Count	Bussines Type	Vountary Deposits	Loan	Class
20	Saving & Loan	774000	4500000	1
14	Food & Beverage	389000	6000000	1
15	Saving & Loan	390000	4000000	1
23	Saving & Loan	38000	1314000	0
14	Saving & Loan	538000	1000000	1
15	Commerce	318000	4000000	1
15	Saving & Loan	414000	5000000	1
20	Commerce	47000	2441000	0
9	Food & Beverage	1087000	3500000	1
17	Commerce	417000	1500000	1
10	Textile	715000	3500000	1
19	Commerce	978000	6000000	1
13	Accesories	328000	4000000	1
18	Commerce	92000	1500000	1
10	Commerce	0	1161000	0
10	Food & Beverage	0	1073000	0
10	Food & Beverage	6000	-	1
10	Saving & Loan	586000	4900000	1
9	Sevices Public	0	803000	0
10	Sevices Public	650000	2000000	1
12	Saving & Loan	308000	1500000	1
5	Food & Beverage	661000	2500000	1
14	Food & Beverage	191000	2949000	0
10	Food & Beverage	673000	5400000	1
10	Commerce	95000	500000	1

Keterangan dimana  
0 = Kredit macet  
1 = Kredit lancar

### Modeling Mesurment

Model yang dihasilkna dari data training akan diuji, menggunakan confusion matrix, bentuk confusion matrix dapat dilihat pada tabel berikut.

Tabel 1. AUC Model Measurment

Correct Labels			
	Positive	Negative	
Positive	True Positive (TP) Flase Negative (FN)	False Positive (FP)	
Negative		True Negative (TN)	

Sumber: Tripathy [9]

Kriteria yang didapat dari *confusion matrix* untuk mengevaluasi kinerja klasifikasi adalah *sensitifity* (Sens), *specificity* (Spec), *positive predictive value* (PPV), *negative predictive value* (NPV), dan *accuracy* (Acc) (Ohsaki et al., 2017). Berikut perhitungannya:

$$\text{Sens} = \frac{TP}{TP+FN} \text{ Equation (1)}$$

$$\text{Spec} = \frac{TN}{TN+FP} \text{ Equation (2)}$$

$$\text{PPV} = \frac{TP}{TP+FP} \text{ Equation (3)}$$

$$\text{NPV} = \frac{TN}{TN+FN} \text{ Equation (4)}$$

$$\text{Acc} = \frac{TP+TN}{TP+FP+TN+FN} \text{ Equation (5)}$$

Kemudian untuk menggambarkan akurasi diskriminan dari suatu pengujian dapat digambarkan dengan Area Under Curva [5], berikut klasifikasi menurut Area Under Curva (AUC)

0.90 – 1.00 = Sangat Baik

0.80 – 0.90 = Baik

0.70 – 0.80 = Cukup Baik

0.60 – 0.70 = Buruk

0.50 – 0.60 = Galat

## RESULT AND DISCUSION

### Model Measurment

K-NN model pada penelitian ini menggunakan K=1, K=3, K=5, K=7, dan K=9.

Nilai K=1, memiliki nilai akurasi 79,45 %, dengan jumlah *True Positive* sebanyak 42 data, *False Positive* sebanyak 8 data, *True Negative* sebanyak 16 data, dan *False Negative* sebanyak 7 data, dengan nilai AUC 0,768

Table 2. TF, PF Rate K=1

	Positive	Negative
Positive	42	8
Negative	7	16

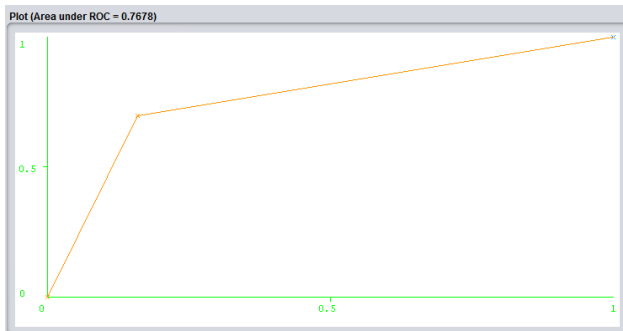


Figure 1. AUC K=1

Nilai K=3 Didapat tingkat akurasi sebesar 72,60%, dengan jumlah *True Positive* sebanyak 40 data, *False Positive* sebanyak 10 data, *True Negative* sebanyak 13 data, dan *False Negative* sebanyak 10 data, dengan nilai AUC 0,748

Table 3. TF, PF Rate K=3

	Positive	Negative
Positive	40	10
Negative	10	13

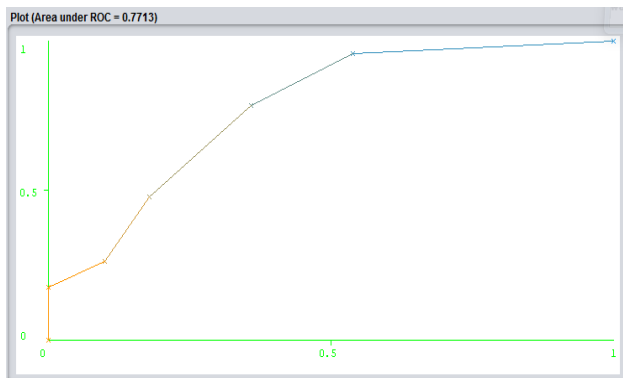


Figure 2. AUC K=3

Nilai K=5 Didapat tingkat akurasi sebesar 71,23%, dengan jumlah *True Positive* sebanyak 41 data, *False Positive* sebanyak

9 data, *True Negative* sebanyak 12 data, dan *False Negative* sebanyak 11 data, dengan nilai AUC 0,771

Table 4. TF, PF Rate K=5

	Positive	Negative
Positive	41	9
Negative	11	12

Dalam grafik dapat dilihat pada gambar AUC di bawah ini:

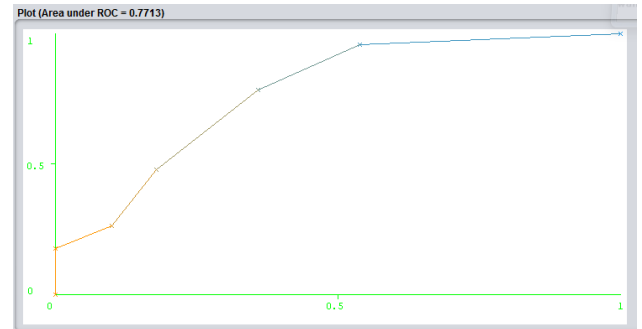


Figure 3. AUC K=5

Nilai K=7 Didapat tingkat akurasi sebesar 72,60%, dengan jumlah *True Positive* sebanyak 41 data, *False Positive* sebanyak 9 data, *True Negative* sebanyak 12 data, dan *False Negative* sebanyak 11 data, dengan nilai AUC 0,804

Table 5. TF, PF Rate K=7

	Positive	Negative
Positive	41	9
Negative	11	12

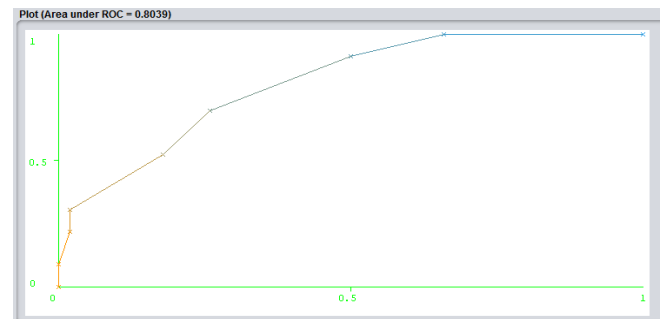


Figure 4 AUC K=7

Nilai K=9 Didapat tingkat akurasi sebesar 72,60%, dengan jumlah *True Positive* sebanyak 44 data, *False Positive* sebanyak 6 data, *True Negative* sebanyak 14 data, dan *False Negative* sebanyak 9 data, dengan nilai AUC 0,811.

Table 6. TF, PF Rate K=9

	Positive	Negative
Positive	44	6
Negative	14	9

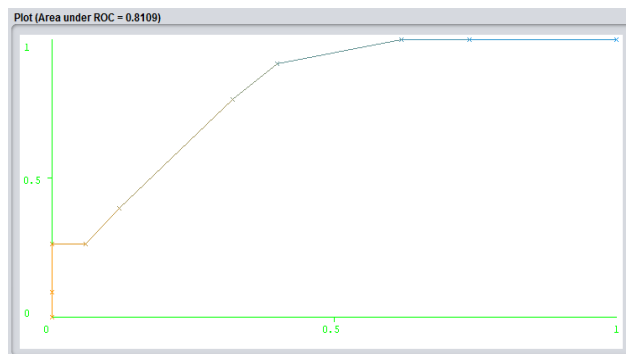


Figure 5. AUC K=9

## CONCLUSIONS

Dari penelitian ini kami mengemukakan ada dua hal penting dalam algoritma K-NN diantaranya Nilai akurasi terbesar terdapat pada nilai K=1 dengan nilai akurasi 79,45%, dengan rata-rata akurasi sebesar 73,696% berdasarkan kurva ROC yang terbentuk dapat dilihat nilai AUC terbesar yang didapat adalah 0,804 pada nilai K=5, dan 0,811 pada nilai K=9, sehingga dapat disimpulkan bahwa algoritma *K-Nearest Neighbor* dapat mengklasifikasi kredit dengan baik.

Untuk penelitian lanjutan kami menyarankan pengoptimalan nilai K yang digunakan menggunakan pembobotan kemudian melakukan optimalisasi pemilihan atribut menggunakan algoritma meta heuristic seperti algoritma genetika, mencoba menerapkan algoritma K-NN kedalam dataset yang lebih besar.

## ACKNOWLEDGMENT

Terimakasih banyak kepada rekan-rekan dosen dan mahasiswa Universitas Bina Sarana Informatika Kampus Sukabumi yang terlibat secara langsung maupun tidak dalam penelitian ini.

## REFERENCES

- [1] Komnas HAM, "http://www.komnasham.go.id/," [Online]. Available: <https://www.komnasham.go.id/files/1475231474-uu-nomor-39-tahun-1999-tentang-SH9FVDS.pdf>. [Accessed 14 05 2019].
- [2] S. Adi and W. E, "Klasifikasi Data NAP (Nota Analisis Pembiayaan) untuk Prediksi Tingkat Kemauan Pemberian Kredit (Studi Kasus: Bank Syariah Mandiri Cabang Luwuk Sulawesi Tengah)," *IJJS (Indonesian Journal of Computing and Cybernetics System)*, vol. 9, no. 1, pp. 1-12, 2015.
- [3] H. Leidiyana, "Penerapan Metode K-Nearest Neighbor pada Penentuan Grade Dealer," *Jurnal Pengetahuan dan Teknologi Komputer*, vol. 2, no. 2, pp. 108-112, 2017.
- [4] R. A. Nugroho, Tarno and P. A, "Klasifikasi Pasien Diabetes Mellitus Menggunakan Metode Smooth Support Vector Machine (SSVM)," *Gaussian*, vol. 6, no. 1, pp. 439-448, 2017.

- [5] M. W. Ohsaki, K. K. Matsuda, H. Watanabe and A. Ralescu, "Confusion Matrix Based Kernel Logistic Regression for Imbalanced Data Classification," *IEEE Transaction on Knowledge and Data Engineering*, vol. 29, no. 9, pp. 1806-1819, 2017.
- [6] T. A. Setiawan, R. S. Wahono and A. Syukur, "Intelligent System," *Journal of Intelligent Systems*, vol. 1, no. 2, pp. 76-81, 2015.
- [7] W. Setiawan and F. Damayanti, "Klasifikasi Citra Retina Menggunakan K-Nearest," *Journal of Intelligent System*, vol. 1, no. 1, pp. 1-6, 2016.
- [8] A. Tripathy, A. Agrawal and S. K. Rath, "Classification of Sentimental Reviews Using Machine Learning Techniques," *Procedia Computer Science*, vol. 5, no. 7, pp. 821-829, 2015.
- [9] T. Wahyono and A. D. Cahyono, "Berbasis Komputasional untuk Meningkatkan Kemampuan," *Jurnal Komputer*, vol. 5, no. 1, pp. 11-18, 2015.
- [10] M. I. M. Zul and L. Hakim, "Human Activity Recognition by Using Nearest Neighbor Algorithm from Digital Image," *ICSIIIT*, vol. 1, no. 1, pp. 58-61, 2017.
- [11] D. Astuti and S. Ati, Pengaruh Implementasi Sistem Manajemen Mutu Iso 9001: 2008 Terhadap Peningkatan Kualitas Layanan UPT Perpustakaan Politeknik Negeri Semarang, Semarang: Ilmu Perpustakaan, 2015.
- [12] S. Bahri, "Seleksi Atribut Pada Algoritma C4.5 Menggunakan Genetik Algoritma dan Bagging Untuk Analisa Kelayakan Pemberian Kredit," *KLIK*, vol. 4, no. 2, pp. 174-183, 2017.
- [13] E. Budianita, J. Jasril and L. Handayani, "Implementasi Pengolahan Citra dan Klasifikasi K-Nearest Neighbor Untuk Membangun Aplikasi Pembeda Daging Sapi dan Babi Berbasis Web," *Jurnal Sains dan Teknologi Industri*, vol. 12, no. 2, pp. 242-247, 2015.

## NOMENCLATURE

Meaning of symbols used in the equations and other symbols presented in your article must be presented in this section.

$\rho$	meaning of $\rho$
$E$	meaning of $E$
$J_c$	meaning of $J_c$

## AUTHOR(S) BIOGRAPHY



### **Rusda Wajhillah**

Anak ke-enam dari tujuh bersaudara. Menyelesaikan Pendidikan formal Pasca Sarjana Ilmu Komputer pada tahun 2013. Aktif mengajar di salah satu Universitas di Indonesia. Selain mengajar, juga aktif dalam pembinaan organisasi mahasiswa di lingkungan kampus.



### **Saeful Bahri**

Memiliki minat dibidang programming, Menyelesaikan Pendidikan pasca sarjana dibidang Ilmu Komputer pada tahun 2016. sejak tahun 2013 sampai saat ini masih aktif sebagai pengajar di salah satu perguruan tinggi swasta terkemuka di Indonesia.

## APPENDICES

Only materials supporting your article can be considered for appendices. Appendices must be presented after Authors Biography. Appendices cannot be more than 4 pages. If your appendices have tables, figures or equations, number them in sequence.